# SAFEPOST

Reuse and development of Security Knowledge assets
for International Postal supply chains

# Risk Assessment Service

*Authors:*
Tommy Westman
FOI

# Contents

Risk Assessment Service User Manual ..... 2
Introduction ..... 2
Limitations ..... 2
Installation ..... 3
Operation ..... 3
Training session ..... 3
Configuration ..... 4
Building risk models, a case study – Theft ..... 4
Building risk models, adding sensor input ..... 10
Summary of the configuration tool ..... 10
Adding document formats ..... 10
Threat handling system ..... 11
Sending, receiving and processing pattern PSS for communicating security related information ..... 11

# Risk Assessment Service User Manual

## Introduction

The core purposes of the Risk Assessment Service are to demonstrate possibilities regarding:

- Analyzing information about the parcel that is electronically available from the postal legacy systems (and or other data sources) before the parcel physically arrives at the postal facility, and to use this information to make an initial "light" risk assessment used for prioritizing certain parcels for more thorough examination using sensors or other equipment once they arrive.
- Fusing the information from the initial "light" risk assessment with outputs from on-site sensor equipment once the parcel physically arrives to produce a more complete risk assessment and to be able to produce a virtual "postal security stamp" (a data file), which includes all the available security information about the parcel the system has been made aware about.

The advanced purposes of the Risk Assessment Service are to demonstrate possibilities regarding:

- Communication regarding security related information about parcels (Pattern PSS). *(Extended pilots only)*
- The postal security stamp and the identified risks functions as input to a system for reasoning about handling the identified threats. *(Extended pilots only)*

### Limitations

The system is only able to operate on whatever information it is provided and what sensors are connected. The system to a large extent uses an expert system approach and therefore the knowledge of a domain expert is required. It is a research prototype, not in any way whatsoever a finished end user product, nor will it be. Therefore user friendliness and general usability are not prioritized areas of concern.

Today item information is captured in an electronic format called ITMATT. This is a suitable document to analyze since it is a document following a postal standard, and it is a document actually used by the postal operators. The benefit of using these messages are that the system will be operating on realistic data and it is therefore easier to make a realistic evaluation of the capabilities. The downside of using this document format is that it contains very, very limited information about the parcel. Only the address information about the recipient (and sometimes the sender), the declared value of the goods, and the weight of the parcel is included. There is a field for describing the contents of the parcel, however in the test data we have viewed this field is either left blank (by far the most common) or it is "free text" in a format which makes it very hard to automatically process.

The system is very configurable and flexible; it is quite easy to add more information sources, everything from manual notes to input from other systems and services or other sensors (that are able to adhere to the necessary data communication format). We can show trivial examples on how such input could be added to the risk analysis and to the postal security stamp, however to add more "real input sources" would require that someone needs to provide them and right now in the project there are no such systems available from any of the participating partners. (One example would be

for instance a customs internal system, but that would require some customs authority to be part of the project and be willing to provide that information and make the systems integration format, and they are not.)

FOI are not domain experts of customs work or risk analysis in the postal domain, so we can only demonstrate how the system could be configured by a domain expert possessing the necessary domain knowledge.

## Installation

The system is installed on a separate standalone server (at the demonstrations FOI will bring it on a laptop or run it from a public server on the internet). It is connected to the CPSS using Ethernet and it receives data through adapters in the CPSS and not directly from the legacy system or the sensors. (In the demo in Iceland, this part must likely be simulated since the legacy system integration or the d-tube integration is not available.)

Thus all that is required for installation is simply network access in between the server and the CPSS. To reach the configuration GUI network access must be available for the client and the server. To connect the D-tube, network access in between the D-tube and the server must be available.

## Operation

The system is not directly operated by any end user. It is connected as a service to the CPSS and thus it serves only other systems (such as the dashboard) within the CPSS.

Internally inside the CPSS it is exposed using a REST API used to both provide the system with data as well as retrieving generated risk assessments. It also connects to the AEON bus to subscribe to data arriving through the adapters.

There is a rough minimal GUI to the application with some web forms to input data and retrieve results. These interfaces are however only intended for testing and demonstration purposes and should not otherwise be used.

## Training session

No training will be performed.

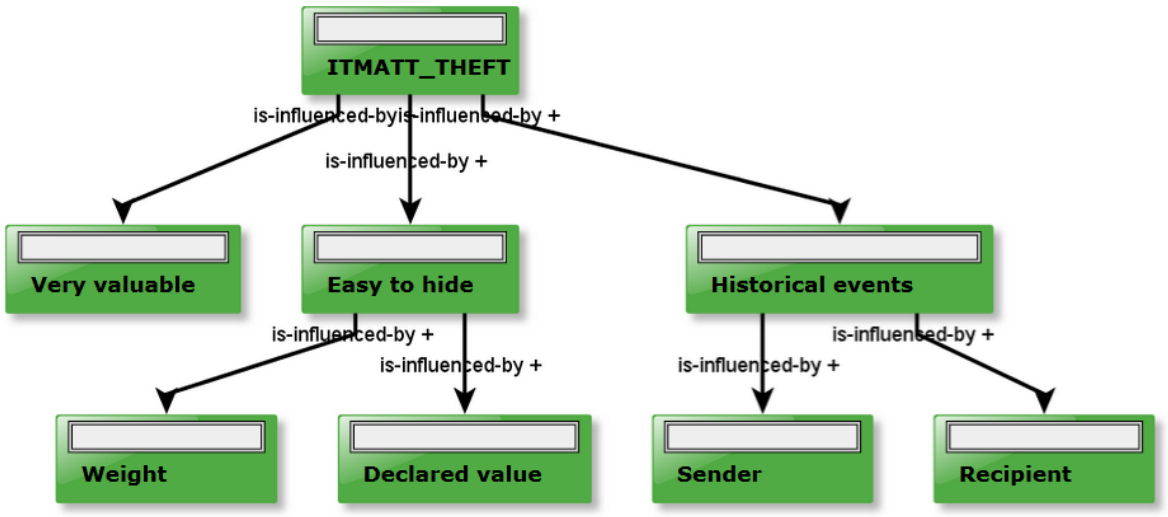
## Configuration

Configuration is the key part to using this system. Configuration is done using a model based approach in a graphical user interface by a designated specialist. It is very much an expert tool, not a system for general use.

### Building risk models, a case study – Theft

This case is chosen for a few reasons

- It is a risk we can analyze using only item information. No sensor would actually provide any valuable input for this type of risk therefore it allows us to focus on one specific input type.
- It is a good example where we can make good use of the limited information available in the ITMATT.
- It is a simple scenario that still is able to show the key parts in constructing a model to handle a specific risk without the model growing too much. Since the scenario is so simple it suits well for serving as an instructive example.
- It includes a few extra things like situations where external services can help out.
- We were able to construct the model without needing too much domain knowledge, just a little bit of reasoning.

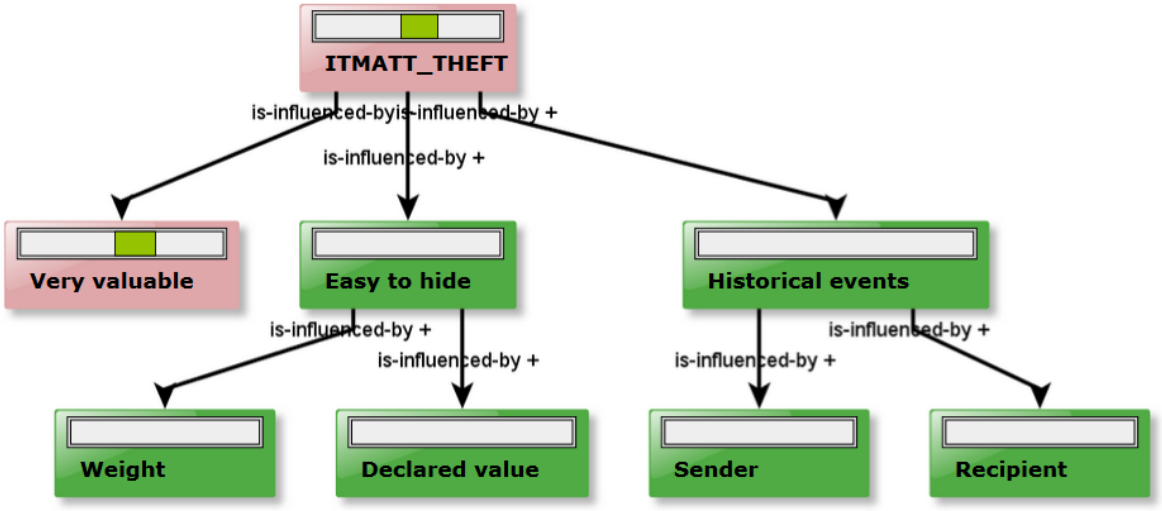



**PICTURE 1: OVERVIEW**

This first picture indicates the finished state of the model. The top node of the model will get its value by different calculations on the values of the nodes below it. The leaf nodes in the model will be set to values defined by input from the document.

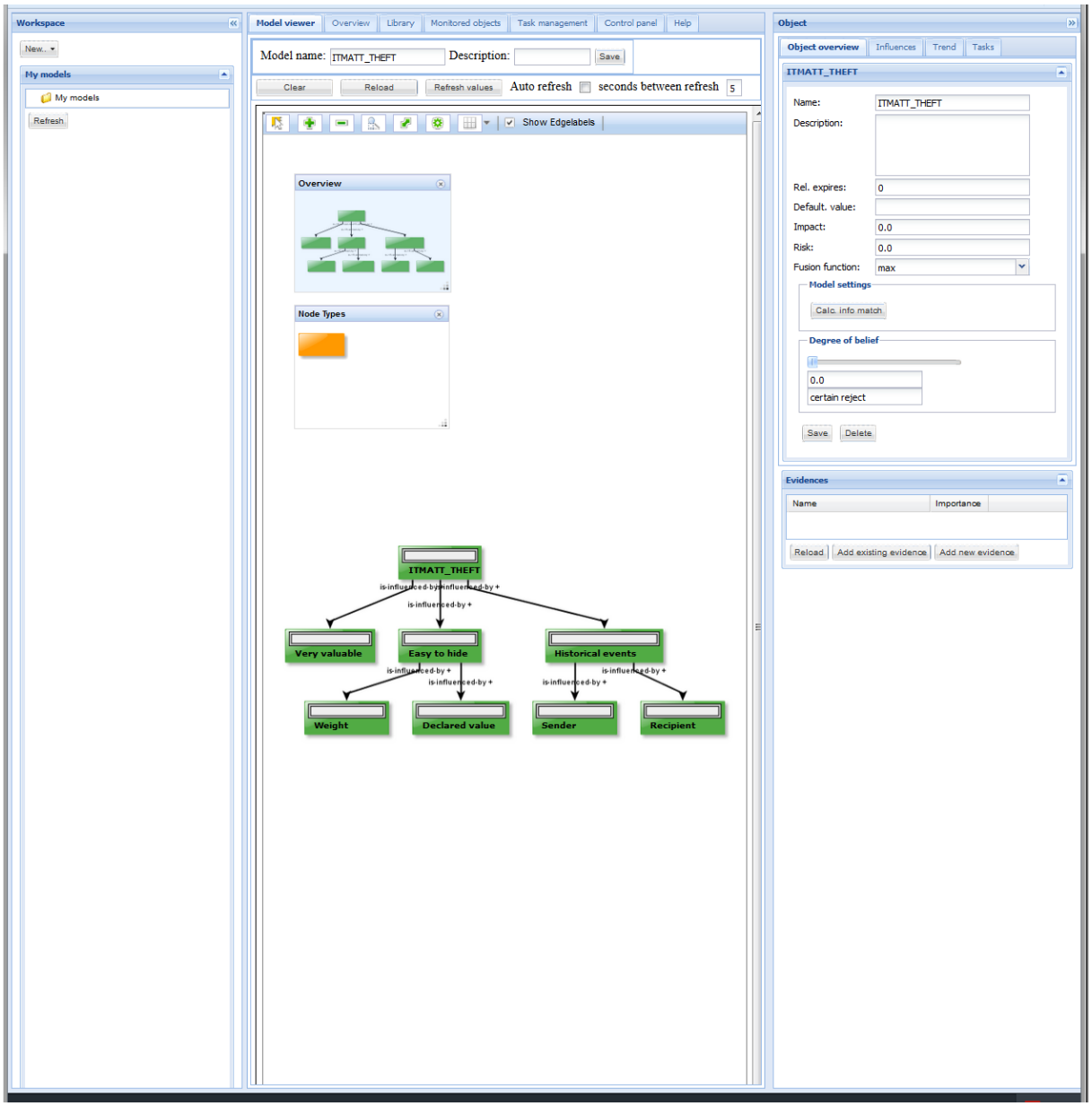
The theft scenario is fairly simple. We will assume there is a risk for theft of a parcel if either of the following circumstances has been met (note that this model does not imply that anyone believes that the parcel actually will be stolen, only that there are reasons to believe that this parcel is more likely to be stolen than other parcels, therefore it could be reasonable to for instance store them in a room with limited access or similar):

1. The parcel is really valuable. What “really valuable” means has to be defined by the domain expert; the goal is to find a level where the extra handling may be justified for insurance reasons, or at a level where the designated storage space is able to host the parcels. (Statistics will help.)
2. The parcel is both valuable and easy to hide (we use the weight of the parcel as an indicator assuming lighter parcels are usually smaller and therefore easier to smuggle out of the facility). The node indicating the parcel value here can be set to a lower value than the other value node.
3. There may have been previous attempts at for instance stealing parcels where a well-known jewelry store is the sender. Perhaps valuables have been sent or it could be as simple as people are trying to find out who has been shopping at that store to rob their homes or similar. The sender could also be some company known to ship very high-end exclusive products such as an art-gallery. These two nodes are used to enter such information. Again, statistics, knowledge about historical incidents, intelligence information or other background knowledge is very useful.



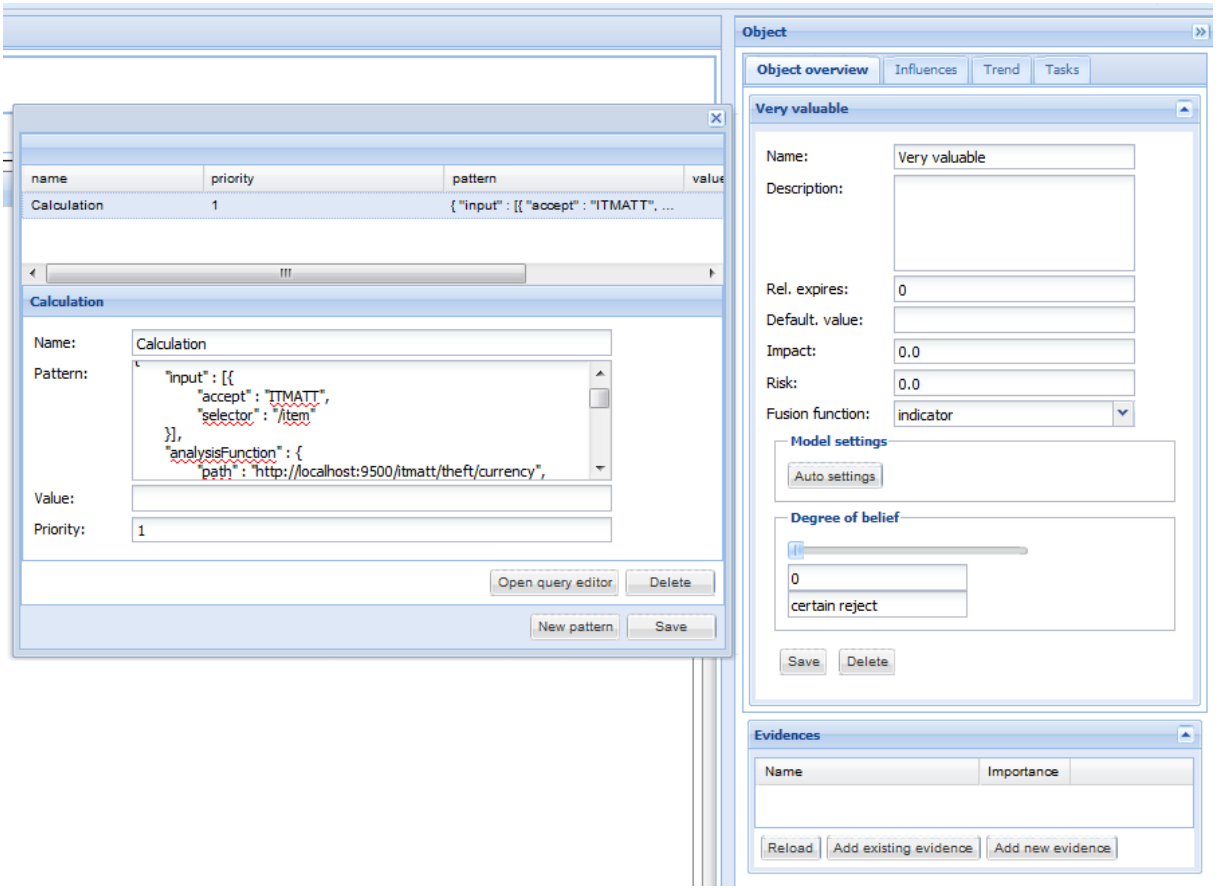

**PICTURE 2: A TRIGGERED MODEL**

The easiest way for this model to trigger is for some leaf node to assume a high value. Depending on which calculation function is set on the root node, different kinds of calculations will be made.




**PICTURE 3: CONFIGURATION TOOL SCREENSHOT**

The models are built in the configuration tool for the Risk Assessment Service. It is a graphical tool where you are able to easily drag nodes onto the canvas and set the properties of the different nodes. In the case of the top node in this model, it has simply been set to use the “max” calculation function (assume the value of the highest value of any direct child node). There are other nodes below using other calculation functions. There are several calculation functions that can be used; a few standard option such as “min”, “max” or “weighted mean”. There are other more advanced options such as using Bayesian networks or custom calculation functions or functions defined in code when the model is calculated.




**PICTURE 4: CALCULATION NODE EXAMPLE**

One example of such a calculation function is found in the leaf nodes that calculate if something is to be regarded as very valuable. When such a node is defined it is possible to give quite exact calculation instructions. They are defined in a well-known standard format called JSON.

```
{
		"input" : [{
				"accept" : "ITMATT",
				"selector" : "/item"
		}],
		"analysisFunction" : {
				"path" : "http://localhost:9500/itmatt/theft/currency",
				"source" : "uri",
				"mimeType" : "application/xml",
				"method" : "POST",
				"takesComparrissonValues" : false
		},
		"compareTo" : {
				"path" : "",
				"source" : "",
				"mimeType" : "",
				"selector" : ""
		}
}
```

**PICTURE 5: CALCULATION INSTRUCTION EXAMPLE**

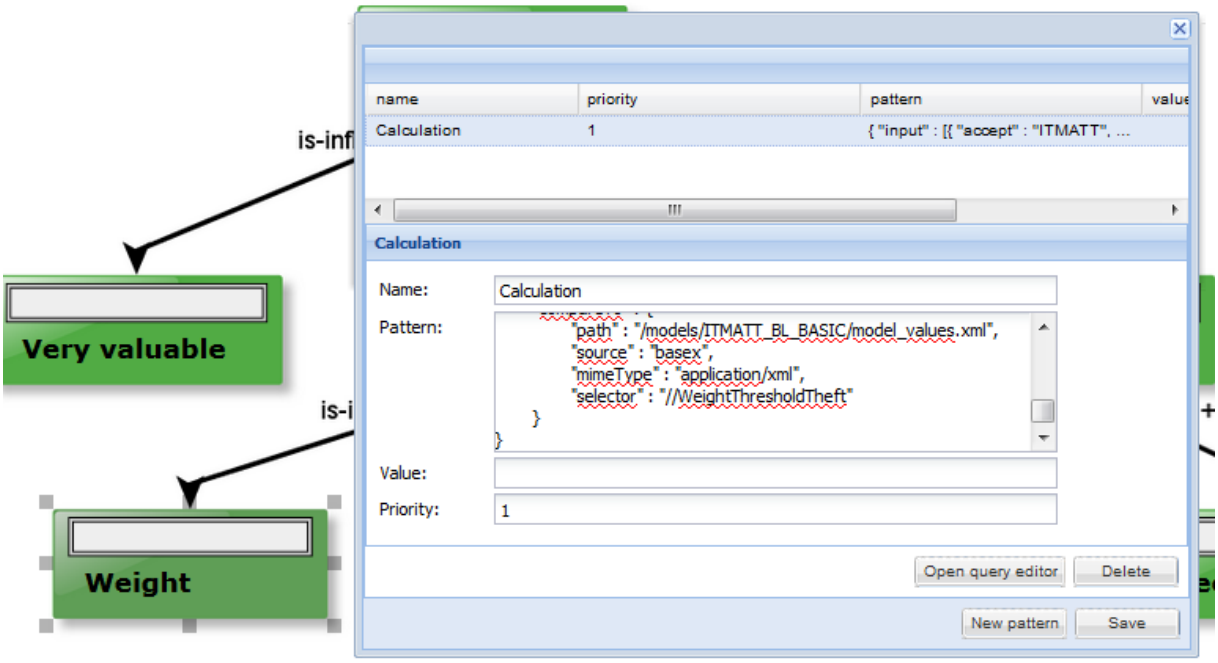
In this case the calculation instruction consists of several parts

| | |
|---|---|
| Input | This section simply tells the risk assessment service where to locate the relevant section of the ITMATT document. In this case it is actually the whole item tag, for reasons described below |
| Analysis function | This is a configuration telling the service what to do with the information. In this case it is using an external service (actually in this case an internal one, but it could just as easily been a third party providing it). Then a few details regarding exactly how to call the service. |
| Compare to | In this case it is not used but it is sometimes used to map the returned values from the external service to certain thresholds. |

The reason for calling an external service in this case is quite simple. When finding the declared value in the customs declaration. It can be expressed in almost any currency. So what the external service simply provides is a currency converter (which the configuration tool does not have). It takes the stated value and converts it to Euro so that it will be simpler to define exact thresholds. Then it returns a risk value depending on what the declared value of the item is. Those thresholds can be as simple as 5000 EUR+ send 1.0, 2500-5000 send 0.8 etc.

The other nodes are using other calculation functions. The sender and recipient nodes also uses an external service (built to parse address lines and extract the name and the location since one of them is not enough, there may be several persons with exactly the same names at different locations, so to be able to identify specific persons both the name and the address is needed). Even though this may seem as a base functionality for a system in the postal domain and perhaps it should be included within the configuration tool, using it as an external service providing the calculations are much more flexible for several reasons.

- It is a modular approach making it possible to improve and upgrade these services independent of the rest of the Risk Assessment Service.
- Third parties can provide these services. It could be that they for instance cooperate with insurance companies to identify these potential targets. Different stakeholders may have different third party providers, a police or customs authority could for instance have a system connected to EUROPOL or similar. The postal operators on the other hand may have customs or police as the third party provider.
- Different internal services can be used by different stakeholders. The postal operators may have a limited need for this type of identification whilst for instance customs or police have access to criminal records and could use it to insert values into the risk calculations without exposing them otherwise.




**PICTURE 6: NODE SETTINGS**

Other types of nodes such as the Weight node do not need any external services. In this case it uses another calculation instruction.

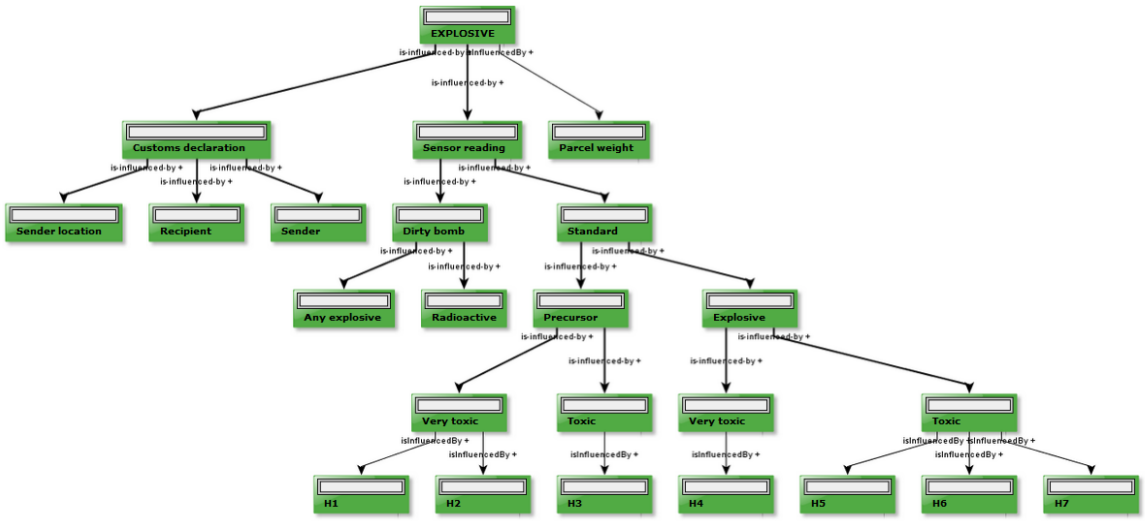
```
{
		"input" : [{
				"accept" : "ITMATT",
				"selector" :
"/interchange/message/msgbody/itmatt_1_2_1/item/declared_gross_weight/text()"
		}],
		"analysisFunction" : {
				"path" : "/models/ra/comparrisson.js",
				"source" : "basex",
				"mimeType" : "application/javascript",
				"method" : "lessThanThreshold",
				"takesComparrissonValues" : true
		},
		"compareTo" : {
				"path" : "/models/ITMATT_BL_BASIC/model_values.xml",
				"source" : "basex",
				"mimeType" : "application/xml",
				"selector" : "//WeightThresholdTheft"
		}
}
```

**PICTURE 7: CALCULATION INSTRUCTION EXAMPLE**

In this case it instead specifies that the specified value of the ITMATT containing the gross weight information should use a simple analysis function specified in a Javascript file to compare it to a list of weights stored in an xml database the application uses to store configuration settings.

## Building risk models, adding sensor input

Not available until the integration with the D-tube and sensors has progressed further. The screenshot below indicates a possible approach. (There is a slight chance that we may be able to show a simulated version already on the Iceland demo.)




**PICTURE 8 : MORE ADVANCED MODEL INCLUDING SENSORS**

## Summary of the configuration tool

We have not in detail described all available options for configuring calculation instructions, nodes settings and model calculation functions, only a few examples. What was important to show is that the configuration of the risk assessment tool is very flexible and can be adapted to very specific needs. It is also important to realize that this system is very much a research prototype. The options will change a lot during development and a production ready system would need to put much more effort into the usability aspects of this tool.

There are many different ways of extending the tools available both by adding connections to external services which both can be data sources, advanced calculation functions, machine learning tools as well as specialist services adapted to specific needs of different actors. If the calculations to be performed are of a fairly trivial nature (similar to blacklists), and therefore it is not needed to develop a specific service for managing them, they are very easy to provide in a simple Javascript file of functions, easily defined and also possible to be changed even during runtime.

# Adding document formats

Adding yet more document formats is as simple as indicating to the system which capabilities each model has regarding which data formats it can analyze and how the system identifies the message being of that format. This is a part that currently does not have a graphical user interface but where there should be one. (The GUI for this has very low priority, since we do not expect more data formats to be added during the life time of the Safepost project.)

Right now any xml based format could be easily added. And then as shown above, configuring the models to use them is as simple as describing where the different information parts are extracted.

A text based format could be added as well as long as it is a structured text format that can be parsed using regular expressions. A binary format could be added (such as image/jpeg or similar) if a corresponding analysis function is constructed. Both of these possibilities is however completely out of scope for this project and will therefore not be described further other than suggestions for possible future extensions.

An example of how such a mapping is defined in the system is the following system definition of ITMATT:

| Field | Definition | Value |
|---|---|---|
| Mime-type | Yes | Application/xml |
| Qualifier | //msgtype/code | ITMATT |
| Identifier | ID/value | |
| Transform | None | |
| Multiple entries | Yes | //item |

The first part “Mime-Type” tells the system which kind of message this is, and if it can be expressed using a standard Mime-Type (some proprietary formats may not have a standard mime type and may require special code to handle); in this case it is an XML message. Knowing what type of message it is it can then parse the rest of the settings knowing how to interpret them.

The qualifier part tells the system how to identify ITMATT messages, in this case by looking for the value “ITMATT” in the path (expressed as) //msgtype/code, knowing it is an XML message the system will use XPath to process it (text messages can for instance be handled with regular expressions).

The identifier tells the system where in the message it can look for an identifier of this very message, which is used to store messages and to link them in the PSS.

Transform is an option that can be used to transform text messages into XML or between XML formats. (XSLT is a well-known technology for this.)

Multiple entries tell the system whether each message contains information about only one item, or if each message possibly contains information about multiple items. In this case since ITMATT messages can contain many items in the same message, it defines how to pick out the individual item entries.

## Threat handling system

Extended pilots only

## Sending, receiving and processing pattern PSS for communicating security related information

Extended pilots only